

Air Quality Assessment by Monitoring PM_{10} and $PM_{2.5}$ Parameters Using Multispectral Satellite Images

Seyed Mohammad Mozafari¹, Ramin Saadi^{1,*}, Sima Neyrizi²

School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran

Faculty of Payambar Azam, Imam Hossein University, Tehran, Iran.

Dep of Geomatics Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

* Corresponding Author: ramin.saadi@yahoo@gmail.com

ABSTRACT

Background and Objectives: Air pollution, particularly particulate matter ($PM_{2.5}$ and PM_{10}), poses significant challenges in large urban areas, leading to severe impacts on public health, ecosystems, and overall quality of life. These issues are especially pronounced in densely populated cities such as Tehran, where air quality management is of utmost importance. Accurate monitoring and forecasting of air quality are essential for developing effective public health and policy strategies. However, the spatial limitations of ground-based air quality monitoring stations prevent comprehensive observation of air quality variations across the entire city. To address these limitations, this study utilized satellite imagery from Landsat-8 and Sentinel-2 to predict particulate matter concentrations, specifically $PM_{2.5}$ and PM_{10} . By combining spectral reflectance data with advanced machine learning methods, the research aims to identify efficient predictive models and determine the most influential spectral bands for estimating particulate matter concentrations.

Materials and Methods: The study began by developing linear regression models using single-band reflectance and multi-band combinations to establish relationships between spectral data and particulate matter concentrations. To capture more complex patterns, nonlinear regression models were also examined. For optimal feature selection, a hybrid Genetic Algorithm-Support Vector Regression (GA-SVR) method was implemented. The Genetic Algorithm (GA) identified the optimal spectral band combinations, while Support Vector Regression (SVR) constructed robust predictive models based on these optimized features. Key evaluation metrics, including the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE), were used to assess and compare model performance. To ensure reliability and generalizability, data were divided into training (70%) and testing (30%) subsets, and cross-validation was applied to validate the models' robustness.

Results and Discussion: The findings revealed that the visible spectrum bands of Landsat-8 and Sentinel-2 showed strong correlations with $PM_{2.5}$ and PM_{10} concentrations. Linear regression models developed using bands 2 and 3 of Landsat-8 and bands 2, 3, and 4 of Sentinel-2 achieved significant correlations in the training

datasets. For Landsat-8, the R^2 values for $PM_{2.5}$ were 0.96% and 9.24% for training and testing datasets, respectively, while Sentinel-2 reached an R^2 of 8.89% for the testing dataset. The RMSE values for Landsat-8 were 0.1 and 0.48 for the training and testing datasets, respectively, while Sentinel-2 demonstrated superior performance with RMSE values of 0.93 and 0.32. These results highlight the effectiveness of Sentinel-2 imagery in predicting particulate matter concentrations.

In the nonlinear regression analysis, power models showed the highest R^2 values among the tested models. The normalized RMSE (NRMSE) values ranged between 0.66 and 0.11, demonstrating greater accuracy than linear models. Although nonlinear models proved more capable of capturing complex relationships, their high computational costs and only marginal accuracy improvements suggest that combining linear models with feature optimization is a more practical approach.

The GA-SVR model yielded the best prediction accuracy, showing that shorter wavelengths play a crucial role in estimating particulate matter concentrations. With optimized feature selection, this model achieved an R^2 close to 0%, underscoring the potential of GA-SVR as a powerful tool for enhancing prediction accuracy in air quality studies.

Conclusion: This study underscores the critical importance of visible spectrum bands in predicting air quality. Sentinel-2 imagery, when combined with the optimal spectral bands identified through the GA-SVR method, demonstrated superior accuracy in estimating $PM_{2.5}$ concentrations. Linear regression models yielded reliable results; however, the integration of feature optimization and advanced machine learning methods significantly enhanced prediction performance. The GA-SVR model achieved remarkable accuracy, with R^2 values as high as 0.96%, underscoring the effectiveness of optimized models for precise and timely air quality monitoring across various spatial scales. These findings highlight the transformative potential of leveraging multispectral satellite imagery alongside machine learning techniques to address the complexities of urban air pollution, offering a robust framework for more informed environmental management and decision-making.

Keywords: Air Pollution, Particulate Matter, Landsat-8, Sentinel-2, AOD

ارزیابی کیفیت هوا با نظارت بر پارامترهای PM_{10} و $PM_{2.5}$ با استفاده از تصاویر ماهواره‌ای چندطیفی

سید محمد مظفری^۱، رامین سعدی^{۲*}، سیما نیری^۳

دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، دانشگاه تهران، تهران، ایران

مرکز علم و فناوری مهندسی ژئوماتیک، دانشکده پیامبر اعظم (ص)، دانشگاه جامع امام حسین (ع)، تهران، ایران

دانشکده ژئوماتیک، دانشگاه ITS، سوربایا، اندونزی

*نویسنده عهده دار مکاتبات: ramin.saadi@yahoo.com

چکیده

سابقه و هدف: آلودگی هوا، به‌ویژه وجود ذرات معلق (PM_{10} و $PM_{2.5}$)، یکی از چالش‌های جدی در شهرهای بزرگ است که پیامدهای مخربی بر سلامت عمومی، اکوسیستم‌ها و کیفیت زندگی دارد. این تأثیرات در مناطق پر جمعیتی مانند تهران به‌وضوح بیشتر احساس می‌شود. از این‌رو، نظارت و پیش‌بینی دقیق کیفیت هوا برای تصمیم‌گیری‌های مؤثر در حوزه سیاست‌گذاری و سلامت عمومی ضروری است. با این حال، محدودیت پوشش مکانی ایستگاه‌های زمینی سنجش کیفیت هوا مانع از ثبت کامل تغییرات کیفیت هوا در سطح شهر می‌شود. در این پژوهش، برای غلبه بر این محدودیت، از تصاویر ماهواره‌ای لندست-۸ و سنتینل-۲ به‌منظور توسعه مدل‌هایی برای پیش‌بینی غلظت ذرات معلق، به‌ویژه $PM_{2.5}$ و PM_{10} ، استفاده شده است. این مطالعه با ترکیب داده‌های بازتاب طیفی ماهواره‌ای و مدل‌های یادگیری ماشین، به شناسایی کارآمدترین مدل‌ها برای پیش‌بینی آلودگی هوا و تعیین باندهای طیفی مؤثر در برآورد غلظت ذرات معلق پرداخته است.

مواد و روش‌ها: در گام نخست، مدل‌های رگرسیون خطی با استفاده از روابط تک‌باندی و ترکیب باندها برای ایجاد ارتباط خطی بین بازتاب طیفی و غلظت ذرات معلق توسعه یافتند. سپس روابط غیرخطی برای مدل‌سازی روابط پیچیده‌تر بازتاب طیفی و ذرات معلق بررسی شدند. برای بهینه‌سازی انتخاب باندهای طیفی، رویکرد ترکیبی GA-SVR به کار گرفته شد؛ به‌طوری‌که الگوریتم ژنتیک (GA) برای شناسایی ترکیب بهینه باندهای طیفی و رگرسیون بردار پشتیبان (SVR) برای ساخت یک مدل پیش‌بینی قوی‌تر روی مجموعه ویژگی‌های بهینه‌شده اعمال شد. معیارهای ارزیابی R^2 ، RMSE و میانگین خطای مطلق (MAE) برای سنجش و مقایسه اثربخشی هر یک از روش‌های مدل‌سازی استفاده شدند. به‌منظور ارزیابی قابلیت تعمیم‌پذیری، داده‌ها به دو بخش آموزشی (۷۰٪) و آزمایشی (۳۰٪) تقسیم شدند. علاوه بر این، برای اطمینان از استحکام مدل، از روش اعتبارسنجی متقابل بهره گرفته شد.

نتایج و بحث: تحلیل‌ها نشان داد که باندهای طیف مرئی در هر دو ماهواره لندست-۸ و سنتینل-۲ ارتباط معناداری با غلظت $PM_{2.5}$ و PM_{10} دارند. مدل‌های رگرسیون خطی با استفاده از باندهای ۱ و ۲ لندست-۸ و باندهای ۲، ۳ و ۴ سنتینل-۲ همبستگی قابل‌توجهی در داده‌های آموزشی ایجاد کردند. مقدار R^2 برای $PM_{2.5}$ در لندست-۸ به ترتیب در داده‌های آموزشی و آزمایشی ۷۰.۵۶٪ و ۶۷.۲۴٪ بود، در حالی که این مقدار در سنتینل-۲ برای داده‌های آزمایشی به ۶۸.۸۹٪ رسید. مقادیر RMSE برای داده‌های آموزشی و آزمایشی لندست-۸ به ترتیب ۷.۰۱ و ۷.۴۸ بود، در حالی که در سنتینل-۲، مقادیر ۶.۹۳ و ۷.۳۲ ثبت شد که نشان‌دهنده عملکرد برتر سنتینل-۲ است.

در مدل‌های رگرسیون غیرخطی، نتایج نشان داد که مدل‌های توانی مقادیر قابل‌توجهی از R^2 را ارائه می‌دهند. در این مدل‌ها، مقادیر NRMSE بین ۰.۰۶۶ تا ۰.۱۱۵ محاسبه شد که نسبت به مدل‌های خطی دقت بالاتری داشتند. اگرچه مدل‌های رگرسیون غیرخطی توانایی بیشتری در مدل‌سازی روابط پیچیده دارند، اما به دلیل بار محاسباتی بالا و افزایش محدود دقت، استفاده از مدل‌های خطی همراه با بهینه‌سازی مبتنی بر یادگیری ماشین، به عنوان رویکردی کارآمدتر توصیه می‌شود.

مدل GA-SVR بالاترین دقت را در پیش‌بینی‌ها داشت و نشان داد که طول‌موج‌های کوتاه‌تر بیشترین توانایی را در پیش‌بینی غلظت ذرات معلق دارند. با بهینه‌سازی انتخاب ویژگی‌ها، این مدل به R^2 نزدیک به ۷۰٪ دست یافت. این نتایج حاکی از توان بالای مدل GA-SVR در بهبود دقت پیش‌بینی است.

نتیجه‌گیری: نتایج نشان داد که باندهای طیف مرئی در مدل‌های پیش‌بینی کیفیت هوا بیشترین تأثیر را دارند. مدل‌سازی با سنتینل-۲ و استفاده از باندهای بهینه شناسایی‌شده توسط روش GA-SVR، بالاترین دقت را در پیش‌بینی $PM_{2.5}$ به دست آورد. تحلیل مقایسه‌ای نشان داد درحالی‌که مدل‌های رگرسیون خطی عملکرد قابل‌قبولی دارند، ترکیب بهینه‌سازی ویژگی‌ها و یادگیری ماشین دقت پیش‌بینی را به‌طور قابل‌توجهی افزایش می‌دهد و اطلاعات ارزشمندی درباره باندهای طیفی مؤثر برای نظارت بر PM ارائه می‌کند. مدل GA-SVR با مقدار R^2 نزدیک به ۷۰٪ نشان‌دهنده قابلیت مدل‌های بهینه شده برای نظارت دقیق و به‌موقع بر کیفیت هوا در مقیاس‌های مکانی مختلف است.

واژه‌های کلیدی: آلودگی هوا، ذرات معلق، لندست-۸، سنتینل-۲، AOD

۱- مقدمه

آلودگی هوا به‌ویژه در کلان‌شهرها و کشورهای در حال توسعه، یکی از مسائل مهمی است که خطرات جدی برای سلامت انسان و محیط‌زیست به همراه دارد (Zou et al., ۲۰۱۹). این پدیده شامل حضور یا ورود بیش از حد مواد مضر مانند گازها، ذرات معلق و مولکول‌های بیولوژیکی در جو است (Glencross et al., ۲۰۲۰; Nasir et al., ۲۰۱۵). پیش‌بینی دقیق کیفیت هوا و ارزیابی آلاینده‌ها گامی اساسی در کاهش این خطرات محسوب می‌شود.

آلودگی هوای شهری سالانه به ۲.۱ تا ۴.۲ میلیون مرگ‌ومیر در سطح جهانی منجر می‌شود و این تعداد در سال ۲۰۱۲ از ۷ میلیون نیز فراتر رفته است (Silva; Brueckner & Pforr, ۲۰۱۰). یکی از آلاینده‌های مهم هوا ذرات معلق جوی (PM) هستند که کیفیت هوا را در مناطق شهری و روستایی تحت تأثیر قرار می‌دهند. ذرات معلق جوی شامل ذرات جامد و مایع با اندازه‌ها و ترکیبات مختلف است که در هوا باقی می‌مانند (Gupta et al., ۲۰۰۶). این ذرات معمولاً به دو دسته $PM_{2.5}$ (ذرات با قطر کمتر

از ۲.۵ میکرون) و PM_{10} (ذرات با قطر کمتر از ۱۰ میکرون) تقسیم می‌شوند که می‌توانند باعث بیماری‌هایی مانند سرطان ریه، بیماری‌های قلبی، آسم و حتی مرگ زودرس شوند (Pope III et al., ۲۰۰۲). علاوه بر این، ذرات معلق آسیب‌های جدی به محیط زیست و بهره‌وری کشاورزی وارد می‌کنند (Nguyen & Tran, ۲۰۱۴).

آلودگی هوا یک چالش زیست‌محیطی فوری در مناطق شهری، به‌ویژه در تهران، به شمار می‌آید، جایی که غلظت ذرات معلق خطرات قابل‌توجهی برای سلامت انسان‌ها ایجاد می‌کند. منابع اصلی آلودگی در تهران شامل وسایل نقلیه، کارخانه‌ها، کارگاه‌های صنعتی، نیروگاه‌ها و پالایشگاه‌ها است که باعث افزایش سطح $PM_{2.5}$ و PM_{10} می‌شوند. این ذرات به عمق سیستم تنفسی نفوذ کرده و منجر به بیماری‌های تنفسی، مشکلات قلبی و حتی مرگ زودرس می‌شوند (Ashmeet Singh et al., ۲۰۲۳; Mozafari et al., ۲۰۱۹). نظارت بر کیفیت هوا برای اجرای تدابیر کنترلی امری ضروری است. ایستگاه‌های نظارت زمینی داده‌های ارزشمندی ارائه می‌دهند، اما پوشش مکانی محدودی دارند و هزینه‌های بالایی برای نگهداری آنها لازم است (Sirsikar & Karemore, ۲۰۱۵).

تصاویر ماهواره‌ای به عنوان روشی جایگزین یا مکمل برای نظارت بر کیفیت هوا مطرح هستند. ماهواره‌های لندست-۸ و سنتینل-۲ به دلیل دقت مکانی و زمانی بالا، امکان ارزیابی دقیق و پیوسته آلاینده‌های PM و منابع آن‌ها را فراهم می‌کنند. ترکیب داده‌های سنجش از دور با اندازه‌گیری‌های زمینی می‌تواند دقت ارزیابی کیفیت هوا را افزایش دهد. مطالعات نشان داده‌اند که ارتباط قوی بین ضخامت نوری هواویزهای استخراج‌شده از تصاویر ماهواره‌ای و اندازه‌گیری‌های PM در ایستگاه‌های زمینی تهران وجود دارد که نشان‌دهنده کارایی سنجش از دور در تخمین سطح آلودگی ذرات است (Mozafari et al., ۲۰۱۹). در مطالعات پیشین، کارایی تصاویر چندطیفی ماهواره‌ای، مانند داده‌های سنجنده MISR بر روی ماهواره Terra و سری ماهواره‌های لندست، در تولید نقشه‌های آلودگی هوا نشان داده شده است (Sameen et al., ۲۰۱۴; Sirsikar & Karemore, ۲۰۱۵).

تصاویر ماهواره‌ای علی‌رغم پوشش مکانی پیوسته و ارائه داده‌های ارزشمند AOD، قادر به اندازه‌گیری مستقیم غلظت‌های PM سطح زمین نیستند (Thorsen et al., ۲۰۲۱). بنابراین، روش‌های مبتنی بر مدل‌سازی برای استخراج غلظت‌های PM سطح زمین از داده‌های AOD ماهواره‌ای همراه با مشاهدات زمینی توسعه یافته‌اند (Mogaraju, ۲۰۲۳). پایه‌گذاری مدل‌سازی آلودگی هوا با استفاده از داده‌های ماهواره‌ای توسط چندین مطالعه انجام شده است (Kaufman et al., ۱۹۹۷). مطالعات بعدی نیز به وابستگی قوی بین غلظت‌های PM و AOT پرداخته‌اند که مبنای تحقیقات بعدی را شکل داده است (Wang & Christopher, ۲۰۰۳). مطالعه (Lim et al., ۲۰۰۴) این رابطه را با استخراج همبستگی تجربی بین AOD و غلظت‌های $PM_{2.5}$ تأیید کرد (Liu et al., ۲۰۰۵). از سنجنده MISR برای استخراج AOD و ایجاد مدل رگرسیون برای $PM_{2.5}$ استفاده کرد. همچنین (Harbula et al., ۲۰۱۰) یک رابطه تجربی بین AOD و PM_{10} با استفاده از داده‌های MODIS ارائه کرد.

اخیراً، کاربردهای سنجش‌ازدور برای مطالعات شهری و تأثیرات آن بر کیفیت هوا بررسی شده‌اند (Mahal et al., ۲۰۲۲; Yu et al., ۲۰۲۲). این مطالعات بر ترکیب داده‌های فتومتر خورشیدی AERONET با مشاهدات ماهواره‌ای تأکید کرده‌اند و بر اهمیت منابع داده یکپارچه برای ارزیابی جامع کیفیت هوا اشاره کرده‌اند. بر اساس تحقیقات پیشین، این مطالعه بر نظارت بر ذرات معلق جوی PM_{10} و $PM_{2.5}$ با استفاده از داده‌های ماهواره‌های لندست-۸ و سنتینل-۲ متمرکز است. ماهواره لندست-۸ هر ۱۶ روز یک بار تصاویر با دقت بالا (۳۰ متر) ثبت می‌کند، در حالی که سنتینل-۲ پوشش بیشتری ارائه می‌دهد و هر ۵ روز یک بار

تصاویر با دقت بالا (در بازه‌های ۱۰ تا ۶۰ متر در باندهای مختلف طیفی) از سطح زمین را ثبت می‌کند. دسترسی رایگان به داده‌های این ماهواره‌ها آن‌ها را به ابزاری کارآمد برای پایش زیست‌محیطی، به ویژه ارزیابی و مدیریت کیفیت هوا تبدیل کرده است. هر دو ماهواره لندست-۸ و سنتینل-۲ پوشش طیفی گسترده‌ای، به‌ویژه در باندهای مرئی، NIR و SWIR، ارائه می‌دهند که برای تجزیه و تحلیل داده‌های کیفیت هوا مناسب هستند.

مطالعات متعددی به بررسی رابطه بین داده‌های ماهواره‌ای و غلظت ذرات معلق هوا پرداخته‌اند. برای مثال، (Saraswat et al. ۲۰۱۷) از تصاویر لندست-۸ جهت تخمین غلظت PM_{10} در شهر دهلی استفاده کردند و با بهره‌گیری از رگرسیون چندگانه، همبستگی مناسبی با داده‌های زمینی گزارش نمودند. همچنین (Gupta & Christopher ۲۰۰۹) ارتباط بین AOD ماهواره MODIS و $PM_{2.5}$ را بررسی کرده‌اند و نتایج آن‌ها بیانگر امکان تخمین غلظت ذرات معلق از تصاویر ماهواره‌ای است. در مطالعه‌ای دیگر، (Kumar et al. ۲۰۲۰) از روش‌های یادگیری ماشین برای تخمین PM با دقت بالا استفاده نمودند.

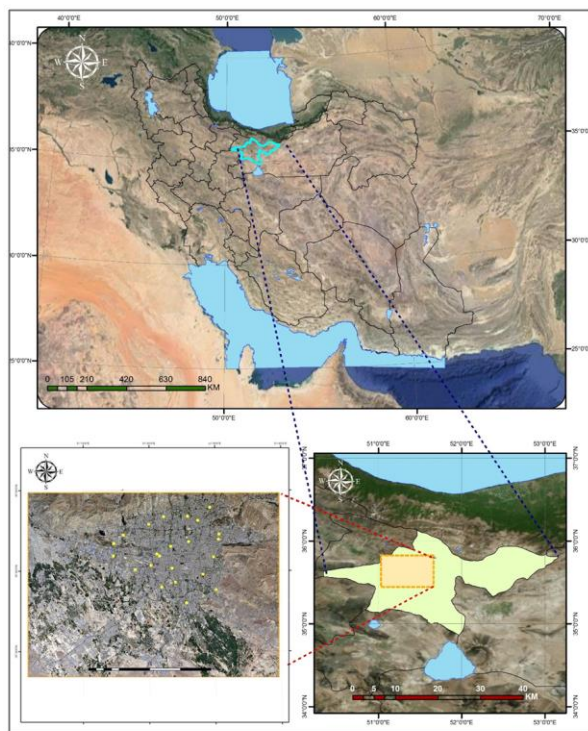
با وجود این پیشینه، اکثر مطالعات پیشین به یکی از موارد زیر محدود بوده‌اند: استفاده از داده‌های با دقت مکانی پایین‌تر مانند MODIS، تمرکز بر تنها یک سنجنده (مانند لندست)، یا بهره‌گیری از مدل‌های آماری ساده. نوآوری این پژوهش در سه بعد اساسی است:

۱. استفاده هم‌زمان از تصاویر دو سنجنده لندست-۸ و سنتینل-۲ با دقت مکانی بالا،
۲. استفاده از شاخص بازتاب اتمسفری (TOA-BOA) به‌عنوان پارامتر تجربی جایگزین AOD،
۳. پیاده‌سازی مدل‌های یادگیری ماشین ترکیبی مانند GA-SVR برای بهبود دقت تخمین غلظت PM در سطح درون‌شهری.

۲- مواد و روش‌ها

۲-۱- منطقه مورد مطالعه

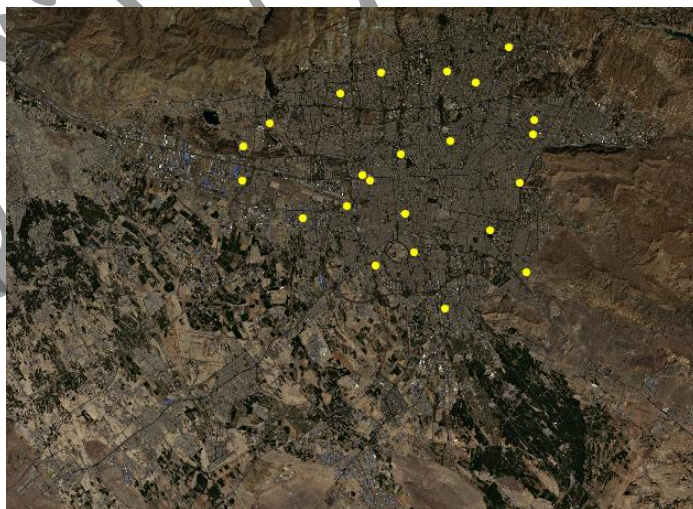
تهران، پایتخت ایران، یکی از کلان‌شهرهای آسیا به شمار می‌آید. این شهر در عرض جغرافیایی میانه کره زمین در ۳۵.۶۸۹۲ درجه شمالی و ۵۱.۳۸۹۰ درجه شرقی قرار دارد (Ganjehi et al., ۲۰۱۷). تهران بزرگ‌ترین شهر ایران است و جمعیتی بیش از ۹ میلیون نفر دارد. بیشتر صنایع ایران در تهران متمرکز شده‌اند و طبق استاندارد سازمان بهداشت جهان (WHO)، تعداد خودروها در تهران تقریباً شش برابر سطح مجاز است. به گزارش شرکت کنترل کیفیت هوای تهران، میانگین تعداد روزهای دارای آلودگی هوا در تهران به ۶۱ روز در سال می‌رسد؛ بنابراین، می‌توان تهران را به‌عنوان یکی از شهرهای آلوده آسیا معرفی کرد. شکل ۱ شهرستان تهران را به‌عنوان منطقه مورد مطالعه نشان می‌دهد که در آن نقاط زرد مکان ایستگاه‌های کنترل کیفیت هوای زمینی را مشخص می‌کنند.



شکل ۱. محدوده مورد مطالعه

۲-۲- داده های میدانی

داده های زمینی در این مطالعه شامل اطلاعات پنج ساله سازمان هواشناسی و شرکت کنترل کیفیت هوای تهران است که به صورت ساعتی برای ۲۳ ایستگاه از سازمان مربوطه دریافت شده است (شکل ۲).



شکل ۲. تصویر توزیع ایستگاه های کنترل کیفیت هوای تهران

داده‌های شرکت کنترل کیفیت هوا مربوط به غلظت آلاینده‌ها می‌باشند که دربردارنده چندین شاخص شامل $PM_{2.5}$ و PM_{10} است و داده‌های هواشناسی نیز شامل سرعت و جهت باد، دمای هوا و فاصله دید افقی می‌باشد. جداول ۱ و ۲ نمونه‌ای از داده‌های زمینی در یک آلوده را نشان می‌دهند.

جدول ۱. شاخص کیفیت هوا (AQI) به تفکیک هر ایستگاه مورخ ۹۷/۱۱/۱۱ ساعت ۱۱

ایستگاه	CO	O ₃	NO ₂	SO ₂	PM ₁₀	PM _{2.5}
میانگین شاخص ایستگاه‌ها	۴۰	۲۲	۷۶	۱۴	۶۵	۸۴
بیشینه شاخص ایستگاه‌ها	۷۰	۵۲	۱۱۳	۲۹	۱۰۳	۱۴۹
اتوبان محلاتی	۲۰	۱۵	۵۵	۵	۵۱	۰
اقدسیه	۴۵	۱۹	۷۵	۰	۴۹	۵۱
پونک	۳۴	۲۰	۱۰۲	۵	۵۳	۶۵
پیروزی	۲۶	۲۶	۵۵	۸	۵۹	۹۴
تربیت مدرس	۵۰	۳۰	۵۱	۸	۶۰	۱۰۹
ستاد بحران	۴۱	۲۴	۷۸	۵	۴۵	۸۴
شادآباد	۵۲	۲۲	۷۳	۱۶	۷۳	۱۲۱

جدول ۲. پارامترهای هواشناسی اخذ شده از ایستگاه‌های زمینی

ایستگاه تهران	
تاریخ	۹۷/۱۱/۱۱
زمان UTC	۰۶
دما	۸°C
فشار	۸۸۴/۵ hpa
نقطه‌ی شبنم	-۳ °C
سرعت باد	۰ m/s
جهت باد	-
دید افقی	۸۰۰۰ متر
بارندگی ۲۴ ساعت گذشته	صفر

تعداد کل داده‌های زمینی مورد استفاده در این پژوهش برابر با ۱۴۴۹ نمونه بوده است که شامل ۹۸۹ نمونه مربوط به تصاویر هم‌زمان لندست-۸ و ۴۶۰ نمونه مربوط به سنتینل-۲ می‌باشد. برای آموزش و آزمون مدل‌ها، از تقسیم‌بندی ۷۰٪ برای آموزش و ۳۰٪ برای آزمون استفاده شد. به این ترتیب، در هر الگوریتم، بخش آموزش و آزمون به صورت تصادفی و مستقل از هم انتخاب گردید تا از بروز overfitting جلوگیری شود.

۲-۳- داده‌های سنجش از دور

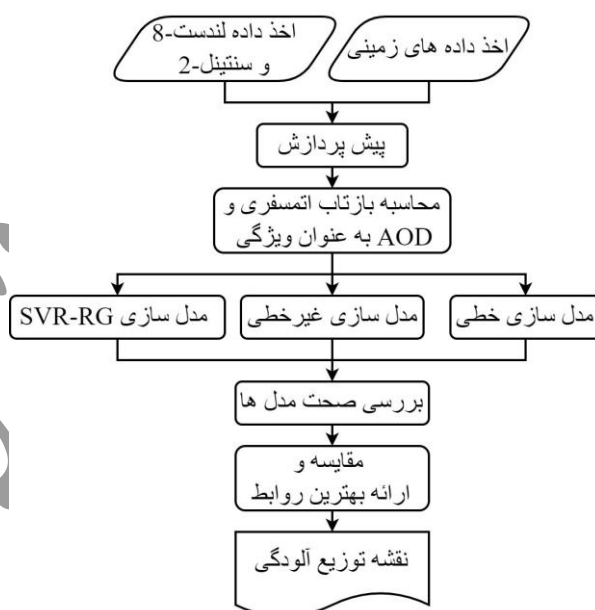
تصاویر سنجش از دور چندطیفی به دلیل پوشش وسیع مکانی، دسترسی آسان و دقت مکانی و طیفی آن‌ها می‌توانند برای ارزیابی کیفیت هوا در مقیاس بزرگ مورد استفاده قرار گیرند. با این حال، همواره بر استفاده از تصاویر با دقت طیفی و مکانی بالاتر برای دستیابی به نتایج بهتر تأکید شده است. در این مطالعه از داده‌های ماهواره لندست-۸ در ۱۱ باند طیفی با دقت مکانی ۳۰ متر و ماهواره سنتینل-۲ در ۱۳ باند طیفی با دقت مکانی ۱۰ تا ۶۰ متر استفاده شده است. مدل‌های آلودگی هوا با استفاده از یک دوره چهار ساله (از ابتدای سال ۲۰۱۵ تا پایان سال ۲۰۱۸ در تهران) از تصاویر سنجنده OLI ماهواره لندست-۸ و برای یک دوره ۱۸ ماهه (از ابتدای سال ۲۰۱۸ تا پایان آگوست ۲۰۱۹ در تهران) از تصاویر سنتینل-۲ توسعه یافته‌اند. جدول ۳ تاریخ تصاویر استفاده شده در این مطالعه را نشان می‌دهد.

جدول ۳. تاریخ تصاویر موجود در منطقه مورد مطالعه بین ژانویه ۲۰۱۵ تا اوت ۲۰۱۹

سنتینل-۲	لندست-۸	
۸ ژانویه ۲۰۱۸	۱۱ نوامبر ۲۰۱۶	۹ ژانویه ۲۰۱۵
۱۴ مارس ۲۰۱۸	۲۷ نوامبر ۲۰۱۶	۱۵ آوریل ۲۰۱۵
۲۴ مارس ۲۰۱۸	۳۰ ژانویه ۲۰۱۷	۵ مه ۲۰۱۵
۳ آوریل ۲۰۱۸	۲۰ آوریل ۲۰۱۷	۲۰ ژوئن ۲۰۱۵
۳ مه ۲۰۱۸	۶ مه ۲۰۱۷	۱۸ ژوئن ۲۰۱۵
۱۲ ژوئن ۲۰۱۸	۲۲ مه ۲۰۱۷	۲۰ ژوئیه ۲۰۱۵
۲۷ ژوئن ۲۰۱۸	۷ ژوئن ۲۰۱۷	۵ اوت ۲۰۱۵

۲۱ اوت ۲۰۱۵	۲۵ ژوئیه ۲۰۱۷	۱۷ ژوئیه ۲۰۱۸
۰۶ سپتامبر ۲۰۱۵	۱۰ اوت ۲۰۱۷	۱۱ اوت ۲۰۱۸
۲۴ اکتبر ۲۰۱۵	۲۶ اوت ۲۰۱۷	۰۵ سپتامبر ۲۰۱۸
۲۵ نوامبر ۲۰۱۵	۱۱ سپتامبر ۲۰۱۷	۲۵ اکتبر ۲۰۱۸
۲۷ دسامبر ۲۰۱۵	۱۳ اکتبر ۲۰۱۷	۱۴ دسامبر ۲۰۱۸
۱۸ ژانویه ۲۰۱۶	۲۹ اکتبر ۲۰۱۷	۲۴ دسامبر ۲۰۱۸
۱۳ فوریه ۲۰۱۶	۱۴ نوامبر ۲۰۱۷	۲۳ ژانویه ۲۰۱۹
۰۳ مه ۲۰۱۶	۳۰ نوامبر ۲۰۱۷	۰۴ مارس ۲۰۱۹
۰۴ جوئن ۲۰۱۶	۱۶ دسامبر ۲۰۱۷	۱۳ آوریل ۲۰۱۹
۲۰ جوئن ۲۰۱۶	۰۱ ژانویه ۲۰۱۸	۱۳ مه ۲۰۱۹
۰۶ ژوئیه ۲۰۱۶	۰۲ فوریه ۲۰۱۸	۰۲ ژوئن ۲۰۱۹
۲۲ ژوئیه ۲۰۱۶	۳۰ سپتامبر ۲۰۱۸	۱۲ ژوئیه ۲۰۱۹
۰۷ اوت ۲۰۱۶	۱۶ اکتبر ۲۰۱۸	۰۱ اوت ۲۰۱۹
۲۳ اوت ۲۰۱۶	۰۱ نوامبر ۲۰۱۸	
۲۴ سپتامبر ۲۰۱۶		

در شکل ۳ مراحل روش پیشنهادی در این پژوهش آورده شده است



شکل ۳. فلوچارت مراحل اصلی مطالعه

۲-۴- پیش پردازش تصاویر ماهواره‌ای

تصاویر ماهواره‌ای لندست-۸ مورد استفاده در این تحقیق از سطح ۱ (Level-1) دریافت شده‌اند. نیاز به در اختیار داشتن هر دو نوع بازتاب TOA و BOA برای محاسبه بازتاب اتمسفری، دلیل انتخاب داده‌های سطح ۱ بود. برای تصاویر اخذ شده می‌توان بازتاب سطح بالای جو (TOA) را استخراج کرد. با این حال، محاسبه بازتاب سطح پایین جو (BOA) نیز برای این مطالعه بسیار حائز اهمیت است، بنابراین نیاز است تا تصحیحات اتمسفری بر روی تصویر اعمال شود.

گرچه داده‌های سطح ۲ (Level-۲) لندست-۸ به صورت آماده (ready-to-use) و شامل بازتاب سطح زمین (BOA) در دسترس هستند، اما در این پژوهش ترجیح داده شد که تصاویر سطح ۱ (Level-۱) مورد استفاده قرار گیرند. این انتخاب به دو دلیل انجام شد: نخست، کنترل دقیق بر پارامترهای تصحیح اتمسفری از جمله مدل جو، مدل آئروسول و دید افقی، که از طریق الگوریتم FLAASH قابل تنظیم و متناسب با شرایط منطقه تهران انجام شد؛ دوم، نیاز به استخراج هم‌زمان بازتاب بالایی (TOA) و پایین (BOA) برای محاسبه بازتاب اتمسفری و شاخص‌های تجربی جایگزین AOD. از آنجا که داده‌های سطح ۲ فقط شامل BOA هستند و TOA خام را ارائه نمی‌دهند، استفاده از تصاویر سطح ۱ و انجام تصحیح اتمسفری به صورت مستقل، امکان دسترسی به هر دو نوع بازتاب را فراهم ساخت. گرچه استفاده ترکیبی از TOA سطح ۱ و BOA سطح ۲ نیز از نظر فنی امکان‌پذیر بود، اما برای حفظ انسجام روش‌شناسی و یکپارچگی پارامترهای تصحیح، استفاده از داده‌های سطح ۱ برای هر دو نوع بازتاب ترجیح داده شد. این تصمیم همچنین موجب افزایش شفافیت مراحل پردازش و قابلیت بازتولید نتایج شد.

مهم‌ترین روش‌های تصحیحات اتمسفری برای لندست-۸ شامل الگوریتم‌های QUAC، ATCOR و FLAASH می‌باشد (Loveland & Irons, 2016). در این پژوهش، برای انجام تصحیح اتمسفری تصاویر لندست-۸ از الگوریتم FLAASH در نرم‌افزار ENVI استفاده شد. بر اساس تاریخ تصویربرداری، عرض جغرافیایی منطقه (۳۵ درجه شمالی) و شاخص‌های هواشناسی، مدل اتمسفری از جدول ۴ و مدل آئروسول Urban انتخاب گردید.

مدل اتمسفری بخش بسیار مهم این الگوریتم است. در حالت کلی مدل‌های اتمسفری مطرح در این الگوریتم به صورت زیر می‌باشند:

- Sub-Arctic Winter
- Mid-Latitude Winter
- U.S. Standard
- Sub-Arctic Summer
- Mid-Latitude Summer
- Tropical

برای انتخاب صحیح مدل اتمسفری مناسب داده‌ی اخذ شده با توجه به تاریخ تصویربرداری و عرض جغرافیایی منطقه مورد مطالعه از جدول زیر استفاده شده است:

جدول ۴. مدل‌های اتمسفری تصحیح FLAASH

نوامبر	سپتامبر	جولای	می	مارس	ژانویه	عرض جغرافیایی
SAW	MLW	MLW	SAW	SAW	SAW	۸۰
SAW	MLW	MLW	MLW	SAW	SAW	۷۰

۶۰	MLW	MLW	MLW	SAS	SAS	MLW
۵۰	MLW	MLW	SAS	SAS	SAS	SAS
۴۰	SAS	SAS	SAS	MLS	MLS	SAS
۳۰	MLS	MLS	MLS	T	T	MLS
۲۰	T	T	T	T	T	T
۱۰	T	T	T	T	T	T

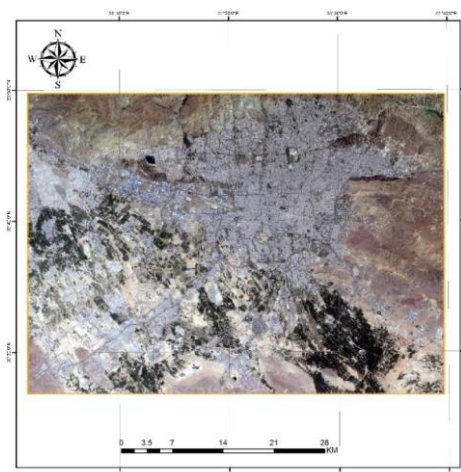
برای مثال با توجه به این که منطقه‌ی مطالعاتی در این تحقیق در عرض جغرافیایی ۳۵ درجه قرار دارد می‌توان دریافت اگر تصویری در ماه ژانویه اخذ شده باشد یکی از بهترین مدل‌های اتمسفری برای آن تصویر مدل Mid-Latitude Summer خواهد بود. برای انتخاب صحیح‌تر و افزایش تأثیرات اتمسفری صحیح در فرآیند تصحیح از جدول پارامترهای هواشناسی زیر استفاده شده است.

جدول ۵. انتخاب مدل اتمسفری بر اساس پارامترهای هواشناسی

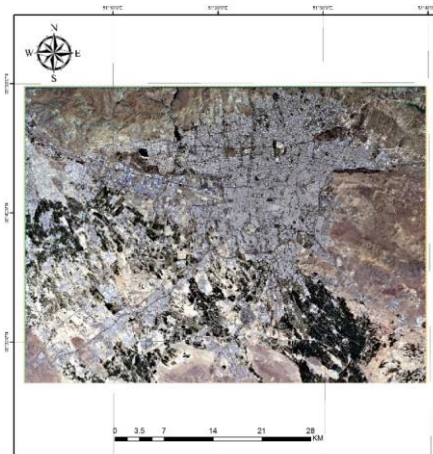
مدل اتمسفری	بخار آب (std atm-cm)	بخار آب (g/cm ^۳)	دمای سطح (درجه سلسیوس)
Winter Sub-Arctic	۵۱۸	۰/۴۲	-۱۶
Winter Mid-Latitude	۱۰۶۰	۰/۸۵	-۱
U.S. Standard	۱۷۶۲	۱/۴۲	۱۵
Sub-Arctic Summer	۲۵۸۹	۲/۰۸	۱۴
Mid-Latitude Summer	۳۶۳۶	۲/۹۲	۲۱
Tropical	۵۱۱۹	۴/۱۱	۲۷

مقدار دید افقی (Visibility) از داده‌های ایستگاه‌های هواشناسی تهران استخراج شد، همچنین ارتفاع متوسط منطقه ۱۲۰۰ متر در نظر گرفته شد. زاویه خورشید و اطلاعات تصویربرداری نیز از فایل metadata فراخوانی و در الگوریتم لحاظ شدند. به منظور انتخاب روش بهینه، مقایسه‌ای بین الگوریتم‌های FLAASH و ATCOR انجام شد که نتایج حاصل از FLAASH دقت بالاتری در تصحیح طیفی نشان داد و به عنوان روش نهایی انتخاب گردید.

تأثیر این تصحیح اتمسفری بر روی تصاویر لندست-۸ در شکل ۴ نشان داده شده است.

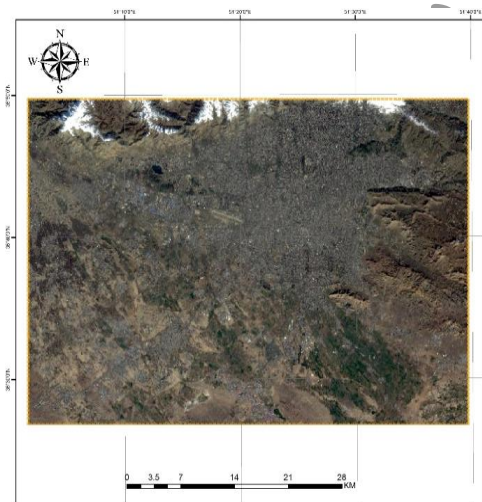


(الف)

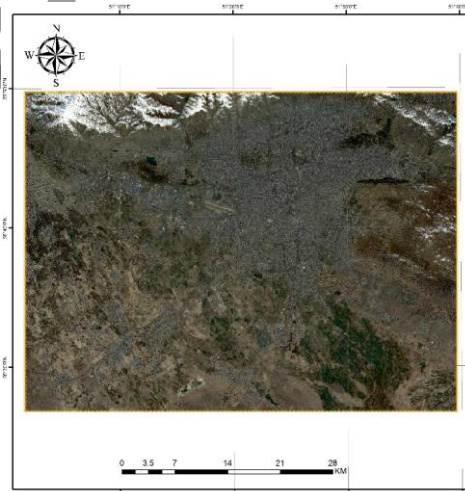


(ب)

شکل ۴. (الف) تصویر ماهواره‌ای لندست-۸ قبل از اصلاح جوی و (ب) همان تصویر پس از اعمال اصلاح جوی FLAASH. در این تحقیق از داده‌های سنتینل-۲ سطح ۱C (L^۱C)، که بازتاب بالای اتمسفر (TOA) را ارائه می‌دهند، استفاده شده است. با این حال، برای تحلیل رگرسیون، داده‌ها باید به سطح ۲A (L^۲A) تبدیل شوند که بازتاب پایین اتمسفر (BOA) را ارائه می‌دهد. این تبدیل با استفاده از افزونه Sen^۲Cor، که پردازشگر مخصوص تولید محصولات سطح A^۲ سنتینل-۲ است، صورت گرفت. تأثیر تصحیح اتمسفری بر تصاویر سنتینل-۲ در شکل ۵ نشان داده شده است.



(الف)



(ب)

شکل ۵. (الف) تصویر ماهواره‌ای سنتینل-۲ قبل از اصلاح جوی و (ب) همان تصویر پس از اعمال اصلاح جوی Sen^۲Cor.

۲-۵- استخراج ویژگی‌های ماهواره‌ای

همانطور که در بخش قبل اشاره شد اولین گام استخراج بازتاب بالای اتمسفر (TOA) از تصاویر ماهواره‌ای است. برای داده‌های سطح ۱C (L^۱C) ماهواره سنتینل-۲، تصاویر ارائه شده به‌طور پیش‌فرض شامل بازتاب بالای اتمسفر (TOA) هستند. اما برای تصاویر لندست-۸، ابتدا باید مقادیر عدد دیجیتال (DN) به بازتاب TOA تبدیل شوند. این تبدیل با استفاده از معادله ارائه شده در رابطه ۱ انجام می‌شود (Sun et al., ۲۰۱۵).

$$\rho_{\lambda'} = M_p \times Q_{cal} + A_p \quad \text{رابطه (۱)}$$

در صورت اعمال تصحیح زاویه خورشید، معادله ۱ به صورت زیر تبدیل می شود:

$$P_{\lambda} = \rho_{\lambda'} / \cos(\theta_{SZ}) = \rho_{\lambda'} / \sin(\theta_{SE}) \quad \text{رابطه (۲)}$$

که در آن، P_{λ} = بازتاب بالای سطح اتمسفر (با تصحیح تأثیر زاویه خورشیدی)، θ_{SZ} = زاویه ارتفاعی محلی خورشید (که در فایل metadata هر تصویر با نام، SUN_ELEVATION موجود است) و θ_{SE} = زاویه زنیته محلی خورشید ($\theta_{SZ} = 90 - \theta_{SE}$) می باشد.

گام بعدی، مشخص کردن تأثیرات اتمسفری بر امواج الکترومغناطیسی رسیده به سنجنده های ماهواره ای است. این تأثیر به صورت تفاوت بین بازتاب اندازه گیری شده توسط سنجنده و بازتاب واقعی از سطح زمین نشان داده می شود. در این مطالعه، این تفاوت به عنوان بازتاب اتمسفری تعریف می شود. بازتاب اتمسفری را می توان به عنوان تفاوت بین بازتاب TOA و بازتاب BOA محاسبه کرد (Saraswat et al., ۲۰۱۷). برای داده های لندست-۸ که هر دو بازتاب TOA و BOA در دسترس هستند، بازتاب اتمسفری را می توان با استفاده از معادله ۳ محاسبه کرد.

$$R(\lambda) = TOA(\lambda) - BOA(\lambda) \quad \text{رابطه (۳)}$$

برای تصاویر سنتینل-۲، همان طور که پیش تر ذکر شد، بازتاب TOA به طور مستقیم در محصول Level-۱C (L1C) ارائه می شود. از سوی دیگر، بازتاب BOA پس از اعمال تصحیح اتمسفری استخراج می شود. بنابراین، بازتاب اتمسفری برای هر باند سنتینل-۲ می تواند با استفاده از معادله زیر (۴) محاسبه شود:

$$R(\lambda) = 1C - 2A \quad \text{رابطه (۴)}$$

که در آن $R(\lambda)$ بازتاب اتمسفری برای باند λ است. در مرحله بعد، برای هر تصویر (تصویر بازتاب اتمسفری حاصل از معادلات ۳ و ۴)، مقدار $R(\lambda)$ متناظر با ایستگاه های کنترل کیفیت هوا استخراج شد. این مقادیر به عنوان داده های سنجش از دور در تحلیل رگرسیون استفاده می شوند.

شایان ذکر است داده های AOD حاصل از ماهواره MODIS با وجود مزایای فراوان، دارای دقت مکانی حدود ۱۰ کیلومتر هستند که برای تحلیل های دقیق در مقیاس شهری مناسب نیستند. در این مطالعه، از اختلاف بین بازتاب TOA و BOA با هدف استخراج شاخصی تجربی از اثرات جو استفاده شد که با رزولوشن بالاتر تصاویر لندست و سنتینل، امکان مدل سازی دقیق تر الگوهای درون شهری آلودگی هوا را فراهم نمود. این رویکرد با وجود محدودیت هایی در مقایسه با AOD، در شرایط فعلی از منظر دقت مکانی مزیت مهمی محسوب می شود.

۲-۶- تحلیل رگرسیون

برای ایجاد یک رابطه تجربی بین داده های میدانی و داده های سنجش از دور، ابتدا تحلیل هایی با استفاده از رگرسیون خطی تک متغیره انجام شد و سپس رگرسیون خطی چندمتغیره مورد بررسی قرار گرفت. در ادامه، تکنیک های پیشرفته تر رگرسیون شامل رگرسیون غیرخطی چندمتغیره (مدل توانی) و رگرسیون بردار پشتیبان با بهینه سازی الگوریتم ژنتیک (SVR-GA) اجرا گردید و نهایتاً ارزیابی و مقایسه نتایج مدل ها انجام شد.

بر اساس مطالعات پیشین، عمق نوری هواویز (AOD) می‌تواند به طور مؤثری با ذرات معلق (PM) دارای همبستگی باشد و به‌این‌ترتیب، به‌عنوان یک شاخص قابل‌اعتماد برای مدل‌سازی آلودگی هوا مورد‌استفاده قرار گیرد. بنابراین مقدار بازتاب اتمسفری حاصل از معادله ۵ برای این منظور مناسب شناسایی شد (Gassó & Hegg, ۲۰۰۳; Fraser et al., ۱۹۸۴). به دلیل رابطه خطی قوی بین AOD و PM، جایگزینی AOD به‌صورت تابعی از PM امکان استخراج معادلات ۶ و ۷ را فراهم می‌کند. این روش الگوریتم برآورد غلظت PM را با استفاده از یک باند طیفی یا طول‌موج خاص (λ) ساده‌تر می‌کند (Retalis & Sifakis, ۲۰۱۰; Sameen et al., ۲۰۱۴)

$$R(\lambda) = TOA(\lambda) - BOA(\lambda) \quad \text{رابطه (۵)}$$

$$AOD = a_1 \times R(\lambda_1) + a_2 \times R(\lambda_2) + \dots + a_n \quad \text{رابطه (۶)}$$

$$PM = a_1 \times R(\lambda_1) + a_2 \times R(\lambda_2) + \dots + a_n \quad \text{رابطه (۷)}$$

۲-۶-۱- رگرسیون خطی چندگانه

معادله ۷ نشان می‌دهد که داده‌های میدانی ارتباط نزدیکی با بازتاب اتمسفری دارند. در این مرحله، با استفاده از معادله ۷، مدل‌های خطی ساده و چندگانه برای ماهواره‌های لندست-۸ و سنتینل-۲ به دست آمدند. با قرار دادن مقادیر میدانی $PM^{۱۰}$ و $PM^{۲.۵}$ در معادله ۷، ضرایب تحلیل رگرسیون برای ماهواره‌های لندست-۸ و سنتینل-۲ محاسبه شدند. سپس، رابطه تجربی بین باندهای اول تا چهارم لندست-۸ و ذرات معلق و باندهای دوم تا چهارم سنتینل-۲ و ذرات معلق استخراج شد.

۲-۶-۲- رگرسیون غیرخطی چندگانه

باتوجه به جدول ۶ و ۷، به نظر می‌رسد که رابطه بین ذرات معلق ($PM^{۱۰}$ و $PM^{۲.۵}$) و بازتاب اتمسفر خطی باشد. برای بررسی این موضوع، یک رابطه غیرخطی بین بازتاب اتمسفری باندهایی که بیشترین همبستگی را با ذرات معلق دارند، برقرار شد. معادله ۸ این رابطه را نشان می‌دهد.

$$PM = a_1 \times R(\lambda_i)^{b_1} + a_2 \times R(\lambda_j)^{b_2} + c \quad \text{رابطه (۸)}$$

در رابطه فوق، a_1 و a_2 ضرایب ثابتی هستند که از تحلیل رگرسیون به دست آمده‌اند، درحالی‌که b_1 و b_2 توان‌های ثابتی هستند که برای ایجاد یک رابطه غیرخطی استفاده می‌شوند. ثابت c به‌عنوان یک پارامتر ثابت اضافی برای تعریف رابطه عمل می‌کند. برای ارزیابی تطابق این معادله با داده‌های موجود، مقادیر بازتاب اتمسفری از باندهای اول و دوم ماهواره لندست-۸ و همچنین باندهای دوم و سوم ماهواره سنتینل-۲ در تحلیل رگرسیون همراه با داده‌های زمینی ذرات معلق استفاده شدند (Loveland & Irons, 2016; Sameen et al., 2014).

Irons, 2016; Sameen et al., 2014)

۲-۶-۳- روش مبتنی بر هسته (Kernel-based method)

از روش SVM برای حل مسائل شناسایی الگو استفاده شده است (Saitta, ۱۹۹۵). در این روش، داده‌ها به یک فضای ورودی با ابعاد بالاتر نگاشته می‌شوند و یک ابرصفحه جداسازی بهینه در این فضا ساخته می‌شود (Huang et al., ۲۰۱۷). برای به کار بردن این روش در اهداف طبقه‌بندی، بیشتر محققان از Support Vector Classification (SVC) استفاده می‌کنند؛ و برای کاربردهای

رگرسیون ، معمولاً از support vector regression (SVR) استفاده می‌شود (Chang & Lin, 2011). در این تحقیق، از روش-ε SVR استفاده شده است.

۲-۷- GA-SVR برای انتخاب ویژگی

فضای ویژگی‌های با ابعاد بالا به دلیل افزایش زمان پردازش و نیاز به تحلیل دقیق، چالش‌های قابل توجهی در اهداف طبقه‌بندی و مدل‌سازی ایجاد می‌کند (Uğuz, 2011; Uysal & Gunal, 2012). بنابراین، انتخاب ویژگی‌های بهینه برای بهبود عملکرد ضروری است. الگوریتم ژنتیک (GA)، یک تکنیک بهینه‌سازی الهام گرفته از مفهوم تکامل است که روشی قدرتمند برای انتخاب ویژگی‌ها و کاهش ابعاد است (Zhong, ۲۰۱۳).

در این مطالعه، استفاده از تمام باندهای طیفی از ماهواره‌های لندست-۸ و سنتینل-۲ برای مدل‌سازی آلودگی هوا غیرعملی و ناکارآمد بود. برای رفع این مشکل، الگوریتم ژنتیک (GA) به کار گرفته شد تا مجموعه بهینه‌ای از باندها برای مدل‌سازی آلودگی هوا شناسایی شود. سپس باندهای منتخب به‌عنوان ورودی برای الگوریتم رگرسیون بردار پشتیبان (SVR) استفاده شدند تا دقت و عملکرد مدل بهبود یابد.

۳- نتایج و بحث

۳-۱- ویژگی‌های ماهواره

مطالعات قبلی نشان داده‌اند که باندهای طیفی ماهواره‌ای در محدوده مرئی به ذرات معلق حساس‌تر هستند. یافته‌های جداول ۲ و ۳، همراه با نتایج تحلیل GA-SVR، تأیید می‌کنند که ذرات معلق بیشترین تأثیر را بر روی باندهای مرئی و طول‌موج‌های کوتاه دارند. این نتایج با نتایج تحقیقاتی قبلی هم‌راستا است؛ بنابراین، برای مدل‌سازی آلودگی هوا، باندهای اول تا چهارم ماهواره لندست-۸ و باندهای دوم تا چهارم ماهواره سنتینل-۲ به‌عنوان مؤثرترین ورودی‌ها انتخاب شدند.

جدول ۶ این روابط را برای ماهواره لندست-۸ و جدول ۷ این روابط را برای ماهواره سنتینل-۲ نشان می‌دهد که با چهار پارامتر دقت (R ، RMSE، R^2) ارزیابی شده‌اند. همچنین، در شکل‌های ۶ و ۷، رابطه خطی تصاویر ماهواره‌ای لندست-۸ و سنتینل-۲ با ذرات معلق قابل مشاهده‌اند.

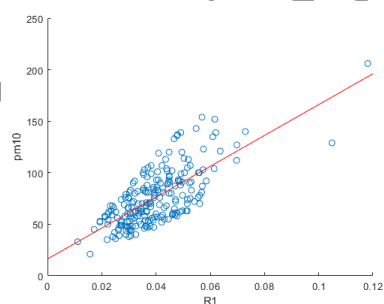
جدول ۶. روابط خطی PM_{10} و $PM_{2.5}$ در جو با استفاده از سنجنده لندست-۸

شماره	مدل خطی	$R^2(\%)$	r	RMSE	NRMSE
۱	$PM_{10} = 1499 \times R(1) + 16,27$	۵۷.۷۶	۰.۷۶	۱۶.۲	۰.۰۸۶
۲	$PM_{2.5} = 616 \times R(1) + 2,27$	۶۸.۸۹	۰.۸۳	۷.۶	۰.۰۷۳
۳	$PM_{10} = 1556 \times R(2) + 24$	۵۱.۸۴	۰.۷۲	۱۷.۳	۰.۰۹۳
۴	$PM_{2.5} = 645 \times R(2) + 5,8$	۶۸.۸۹	۰.۸۳	۸.۱	۰.۰۷۷
۵	$PM_{10} = 1146 \times R(3) + 82$	۳۲.۴۹	۰.۵۷	۲۳	۰.۱۲۴
۶	$PM_{2.5} = 624 \times R(3) + 30$	۵۰.۴۱	۰.۷۱	۱۰	۰.۰۹۶
۷	$PM_{10} = 983 \times R(4) + 88,9$	۲۲.۰۹	۰.۴۷	۲۴.۱	۰.۱۲۹
۸	$PM_{2.5} = 560 \times R(4) + 35$	۴۶.۲۴	۰.۶۸	۱۱	۰.۱۰۵

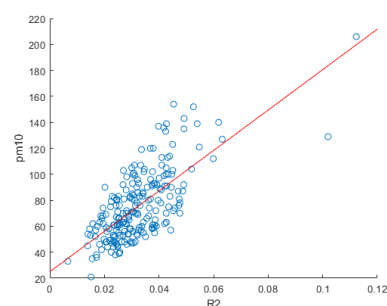
۹	$PM_{10} = 1885 \times R(1) - 486 \times R(2) + 14$	۶۲.۴۱	۰.۷۹	۹.۲	۰.۰۴۹
۱۰	$PM_{2,5} = 80 \times R(1) + 559 \times R(2) + 5,3$	۶۸.۸۹	۰.۸۳	۷.۲	۰.۰۶۹
۱۱	$PM_{10} = 3858 \times R(3) - 2763 \times R(4) + 58$	۴۰.۹۶	۰.۶۴	۲۰.۱۸	۰.۱۰۹
۱۲	$PM_{2,5} = -1013 \times R(3) + 1675 \times R(4) + 22,2$	۵۶.۲۵	۰.۷۵	۹.۸	۰.۰۹۴
۱۳	$PM_{10} = 1575 \times R(1) + 279 \times R(2) - 435 \times R(3) + 1,6$	۶۴.۹۶	۰.۸۰۶	۱۶.۹	۰.۰۹۱
۱۴	$PM_{2,5} = -52 \times R(1) + 781 \times R(2) - 96 \times R(3) + 2,9$	۶۸.۸۹	۰.۸۳	۷.۸	۰.۰۷۵
۱۵	$PM_{10} = 1422 \times R(1) + 424 \times R(2) - 447 \times R(4) - 1$	۶۲.۴۱	۰.۷۹	۱۸.۵	۰.۱۰
۱۶	$PM_{2,5} = -194 \times R(1) + 954 \times R(2) - 126 \times R(4) + 1,7$	۶۸.۸۹	۰.۸۳	۷.۹	۰.۰۷۵
۱۷	$PM_{10} = 1359 \times R(1) + 397 \times R(2) + 457 \times R(3) - 822 \times R(4) - 0,2$	۶۴.۰۰	۰.۸۰	۱۹.۶	۰.۱۰۵
۱۸	$PM_{2,5} = -221 \times R(1) + 1028 \times R(2) + 378 \times R(3) - 435 \times R(4) + 2,5$	۷۰.۵۶	۰.۸۴	۷.۱	۰.۰۶۸

جدول ۷. روابط خطی PM_{10} و $PM_{2,5}$ در جو با استفاده از سنجنده سنتینل-۲

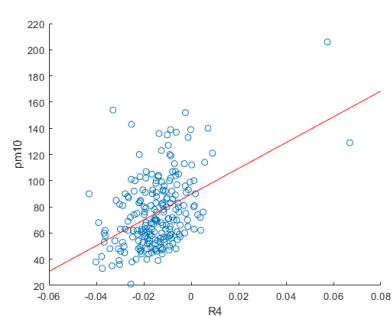
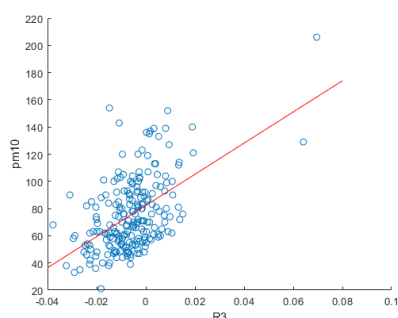
شماره	مدل خطی	$R^2(\%)$	r	RMSE	NRMSE
۱	$PM_{10} = 0,0841 \times R(2) + 15$	۶۲.۴۱	۰.۷۹	۸.۱	۰.۱۱۵
۲	$PM_{2,5} = 0,0452 \times R(2) - 1,46$	۶۴.۹۶	۰.۸۰۶	۷.۲	۰.۱۰۲
۳	$PM_{10} = 0,159 \times R(3) + 31$	۵۱.۸۴	۰.۷۲	۲۰.۶	۰.۱۵۱
۴	$PM_{2,5} = 0,0757 \times R(3) + 8$	۶۷.۲۴	۰.۸۲	۷.۹	۰.۱۱۲
۵	$PM_{10} = 0,0887 \times R(4) + 59$	۳۱.۳۶	۰.۵۶	۲۴	۰.۱۷۶
۶	$PM_{2,5} = 0,0352 \times R(4) + 21$	۳۷.۲۱	۰.۶۱	۱۰.۱	۰.۱۴۴
۷	$PM_{10} = 0,0771 \times R(2) - 0,0140 \times R(3) + 14$	۶۴.۳۲	۰.۸۰۲	۱۲	۰.۱۰۱
۸	$PM_{2,5} = 0,0349 \times R(2) + 0,0101 \times R(3) - 1,1$	۶۷.۲۴	۰.۸۲	۷.۲۷	۰.۱۰۳
۹	$PM_{10} = 0,0654 \times R(2) + 0,0428 \times R(4) + 23$	۳۸.۴۴	۰.۶۲	۱۴.۱	۰.۱۱۹
۱۰	$PM_{2,5} = 0,0532 \times R(2) + 0,0341 \times R(4) + 12$	۴۰.۹۶	۰.۶۴	۱۰.۳	۰.۱۴
۱۱	$PM_{10} = 0,102 \times R(3) - 0,024 \times R(4) + 35$	۳۸.۴۴	۰.۶۲	۱۶.۴	۰.۱۳۸
۱۲	$PM_{2,5} = 0,0497 \times R(3) - 0,009 \times R(4) + 13$	۴۳.۵۹	۰.۶۶	۹.۱۸	۰.۱۳
۱۳	$PM_{10} = 0,074 \times R(2) + 0,015 \times R(3) - 0,007 \times R(4) + 14$	۶۸.۸۹	۰.۸۳	۱۲.۳	۰.۱۰۴
۱۴	$PM_{2,5} = 0,041 \times R(2) + 0,012 \times R(3) - 0,007 \times R(4) - 2,2$	۷۰.۵۶	۰.۸۴	۶.۹۳	۰.۰۹۹



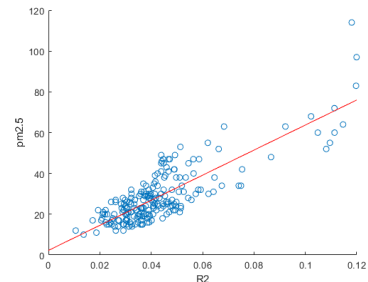
(الف) نمودار خطی بازتاب اتمسفری باند اول لندست-۸ با PM_{10}



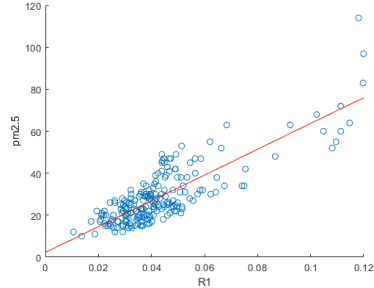
(ب) نمودار خطی بازتاب اتمسفری باند دوم لندست-۸ با PM_{10}



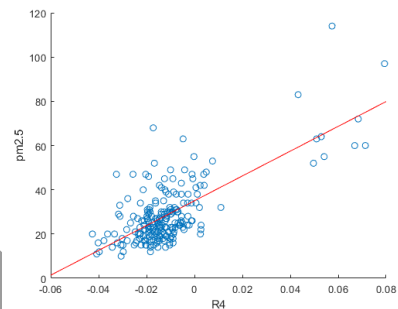
(د) نمودار خطی بازتاب اتمسفری باند چهارم لندست-۸ با $PM_{1.0}$



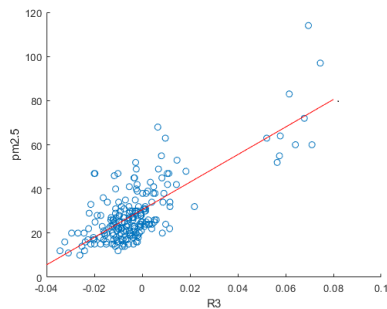
(ج) نمودار خطی بازتاب اتمسفری باند سوم لندست-۸ با $PM_{1.0}$



(و) نمودار خطی بازتاب اتمسفری باند اول لندست-۸ با $PM_{2.5}$



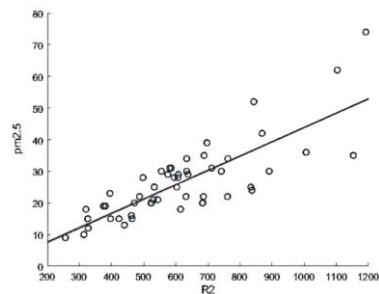
(ه) نمودار خطی بازتاب اتمسفری باند اول لندست-۸ با $PM_{2.5}$



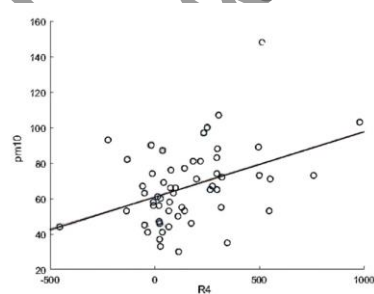
(ح) نمودار خطی بازتاب اتمسفری باند چهارم لندست-۸ با $PM_{2.5}$

شکل ۶. (الف) تا (د) رابطه خطی بین باندهای اول تا چهارم تصاویر لندست-۸ و غلظت‌های $PM_{1.0}$ را به ترتیب نشان می‌دهند. به طور مشابه، (ه) تا (ج) رابطه خطی بین باندهای اول تا چهارم تصاویر لندست-۸ و غلظت‌های $PM_{2.5}$ را نمایش می‌دهند. روابط تجربی دقیق و معیارهای دقت مربوط به این نمودارها در جدول ۶ ارائه شده است.

(ب) نمودار خطی بازتاب اتمسفری باند سوم سنتینل-۲ با $PM_{1.0}$

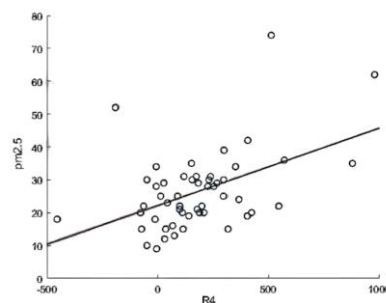
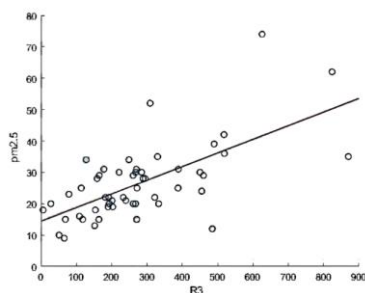


(الف) نمودار خطی بازتاب اتمسفری باند دوم سنتینل-۲ با $PM_{1.0}$



(د) نمودار خطی بازتاب اتمسفری باند دوم سنتینل-۲ با $PM_{2.5}$

(ج) نمودار خطی بازتاب اتمسفری باند چهارم سنتینل-۲ با $PM_{1.0}$



(ه) نمودار خطی بازتاب اتمسفری باند سوم سنتینل-۲ با $PM_{2.5}$

(و) نمودار خطی بازتاب اتمسفری باند چهارم سنتینل-۲ با $PM_{2.5}$

شکل ۷. روابط خطی بین باندهای دوم تا چهارم تصاویر سنتینل-۲ و PM_{10} در نمودارهای (الف) تا (ج) نشان داده شده است. به طور مشابه، نمودارهای (د) تا (و) روابط خطی بین همان باندها و $PM_{2.5}$ را نشان می‌دهند. یافته‌های تجربی، شامل روابط و معیارهای دقت برای این نمودارها، در جدول ۷ به تفصیل بیان شده است.

۲-۳- مدل‌سازی آلودگی هوا

سه روش رگرسیون خطی، رگرسیون غیرخطی و GA-SVR برای مدل‌سازی آلودگی هوا ناشی از ذرات معلق استفاده شدند.

۱-۲-۳- مدل‌سازی خطی

تحلیل اولیه در مدل‌سازی خطی، شامل استخراج روابط خطی برای هر باند انتخابی در طیف الکترومغناطیسی بود که پس از آن روابط چندجمله‌ای بین باندها ایجاد شد. طبق جدول ۶، برای ماهواره لندست-۸، باند اول قوی‌ترین پاسخ را به ذرات معلق نشان می‌دهد، و باند دوم نیز همبستگی قابل توجهی دارد. مدل‌های چندجمله‌ای خطی که از باندهای اول تا سوم استفاده می‌کنند، دقت کمی بیشتر از سایر مدل‌ها نشان دادند. با این حال، به منظور بهینه‌سازی کارایی محاسباتی، استفاده از مدل چندجمله‌ای تنها با باندهای اول و دوم توصیه می‌شود. برای ماهواره سنتینل-۲، همان‌طور که در جدول ۷ مشاهده می‌شود، باند ۲ قوی‌ترین همبستگی را با ذرات معلق دارد. علاوه بر این، مدل چندجمله‌ای که شامل باندهای ۲، ۳ و ۴ است، بالاترین همبستگی را نشان می‌دهد. با این حال، برای ساده‌سازی محاسبات و با صرف نظر از کردن کمی کاهش دقت (حدود ۲٪ همبستگی)، یک مدل چندجمله‌ای خطی با استفاده از باندهای ۲ و ۳ می‌تواند به طور مؤثر به عنوان مدل پیش‌بینی آلودگی هوا عمل کند. جداول ۲ و ۳ روابط خطی بین ذرات معلق و تصاویر ماهواره‌ای لندست-۸ و سنتینل-۲ را به تفصیل نشان می‌دهند.

۲-۲-۳- مدل‌سازی غیرخطی

توسعه مدل‌های غیرخطی برای بررسی جامع‌تر رابطه بین ذرات معلق و بازتاب اتمسفری ثبت‌شده در تصاویر سنجش‌ازدور بسیار مفید هستند. در این مطالعه، رگرسیون غیرخطی (به‌ویژه روابط توان) برای دو باند که بیشترین همبستگی را با ذرات معلق نشان می‌دادند، اعمال شد. یافته‌های این مدل‌ها نشان داد حتی زمانی که از روش‌های غیرخطی برای تجزیه و تحلیل استفاده شود بازتاب اتمسفری عمدتاً رابطه‌ای خطی با داده‌های آلودگی دارد. جداول ۸ و ۹ مدل‌های تجربی تولید شده را همراه با پارامترهای دقت مربوطه ارائه می‌دهند.

جدول ۸. روابط غیرخطی PM_{10} و $PM_{2.5}$ در جو با استفاده از سنجنده لندست-۸

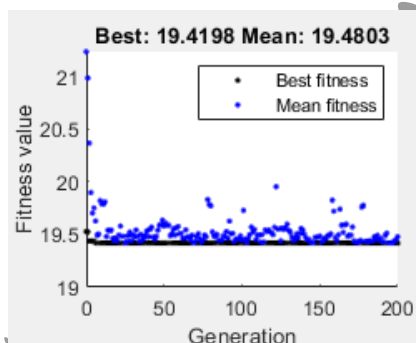
شماره	مدل غیرخطی	$R^2(\%)$	r	RMSE	NRMSE
۱	$PM_{10} = 866 \times R(1)^{1.085} + 755 \times R(2)^{0.897} + 8,1$	۶۰,۸۴	۰,۷۸	۱۹,۱	۰,۱۰۶

جدول ۹. روابط غیرخطی PM_{10} و $PM_{2.5}$ در جو با استفاده از سنجنده سنتینل-۲

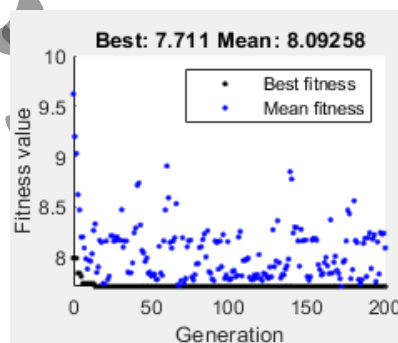
شماره	مدل غیرخطی	$R^2(\%)$	r	RMSE	NRMSE
۱	$PM_{10} = ۰,۰۸۴ \times R(۲) + ۱,۲۴ - ۶,۴۳ \times R(۳) + ۰,۰۴۱ + ۱۱$	۶۵,۶۱	۰,۸۱	۷,۳	۰,۰۶۶
۲	$PM_{2.5} = ۰,۰۰۹۷ \times R(۲) + ۱,۳۳ - ۰,۹ \times R(۳) + ۰,۳۴ + ۳۰$	۶۴,۹۶	۰,۸۰۶	۸,۱	۰,۱۱۵

۳-۲-۳ مدل سازی مبتنی بر هسته (Kernel-based modeling)

استفاده از الگوریتم های یادگیری ماشین برای شناسایی مؤثرترین ویژگی ها از یک مجموعه بزرگ ضروری است. هدف اصلی در این تحقیق، تعیین مناسب ترین باندهای ماهواره ای برای توسعه مدل های آلودگی هوا بر اساس ذرات معلق در جو بود. با استفاده از الگوریتم SVR-GA (رگرسیون بردار پشتیبان با الگوریتم ژنتیک) و استفاده از کمینه سازی RMSE به عنوان تابع هدف، باندهای طیفی بهینه برای هر ماهواره شناسایی شدند. نتایج حاصل از الگوریتم SVR-GA نتایج بخش های قبلی را تأیید کرد و نشان داد که باندهای با طول موج کوتاه تر ارتباط قوی تری با ذرات معلق در جو دارند. به طور خاص، داده های ورودی برای الگوریتم SVR-GA شامل باندهای ۱ تا ۷ از ماهواره لندست-۸ و باندهای ۱ تا ۱۲ از ماهواره سنتینل-۲ بودند. مطابق با انتظارات، الگوریتم باندهای با طول موج کوتاه تر را به عنوان ویژگی های مرتبط تر انتخاب کرد. نتایج این تحلیل ها در شکل های ۵ و ۶ نشان داده شده است، در حالی که شکل های ۸ و ۹ کاربرد الگوریتم GA-SVR را برای ماهواره های لندست-۸ و سنتینل-۲ به ترتیب نشان می دهند.



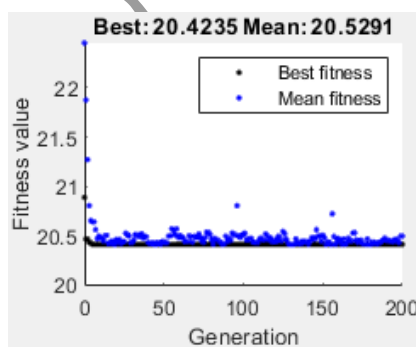
(الف)



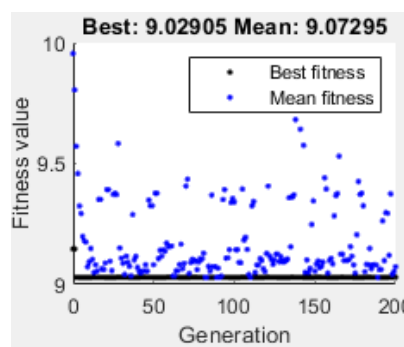
(ب)

شکل ۸. (الف) الگوریتم GA-SVR که بر رابطه تصاویر ماهواره ای لندست-۸ و PM_{10} اعمال شده است و (ب) الگوریتم

GA-SVR که بر رابطه تصاویر ماهواره ای لندست-۸ و $PM_{2.5}$ اعمال شده است



(الف)



(ب)

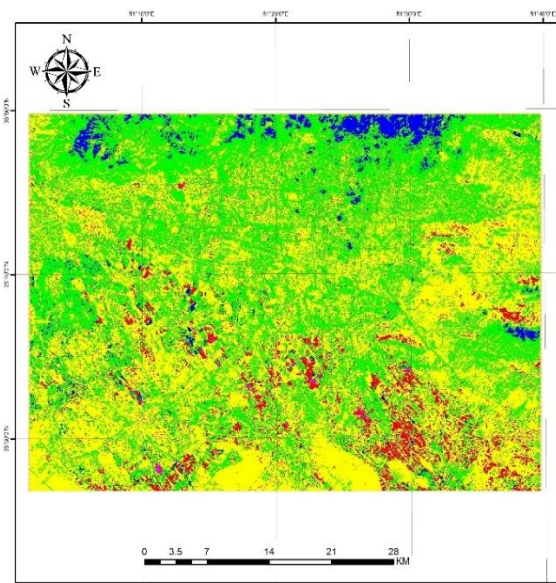
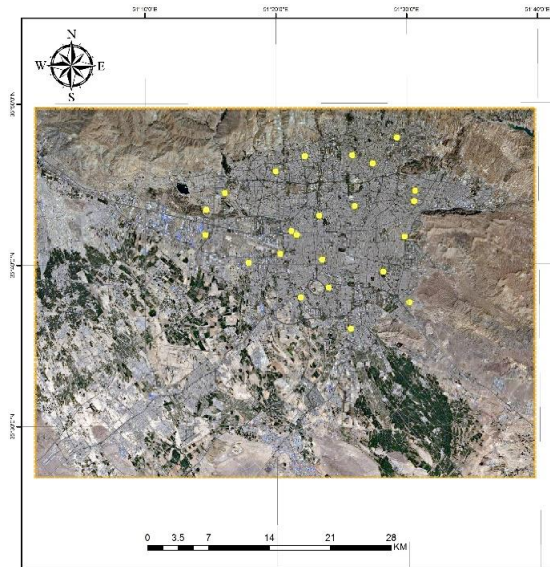
شکل ۹. الف) الگوریتم GA-SVR که بر روی رابطه تصاویر ماهواره‌ای سنتینل-۲A و PM_{10} اعمال شده است و ب) الگوریتم GA-SVR که بر روی رابطه تصاویر ماهواره‌ای سنتینل-۲A و $PM_{2.5}$ اعمال شده است.

۳-۳- پیش بینی نقشه آلودگی هوا

نتایج تحلیل رگرسیون نشان می‌دهد که استفاده از باندهای طول موج کوتاه‌تر برای توسعه مدل‌های پیش‌بینی آلودگی هوا بسیار مؤثر است. به طور خاص، ماهواره سنتینل-۲ بهبود دقت حدود دو درصدی نسبت به ماهواره لندست-۸ ارائه داده است. با این حال، علی‌رغم همبستگی‌های قوی یافت شده در مدل‌های غیرخطی، به دلیل پیچیدگی افزوده و نیاز به محاسبات بیشتر، این مدل‌ها برای برآورد آلودگی هوا توصیه نمی‌شوند. به عنوان مثال، سطوح آلودگی هوا در تهران در تاریخ ۲۹ اکتبر ۲۰۱۷، که یکی از آلوده‌ترین روزهای ثبت شده است، با استفاده از این روش برآورد شد. پس از دریافت تصویر ماهواره‌ای لندست-۸ برای این تاریخ، معادله ردیف نهم جدول ۶ بر داده‌های بازتاب اتمسفری اعمال شد. نقشه آلودگی هوا برای PM_{10} که از این روش حاصل شده، در شکل ۹ نشان داده شده است و به خوبی با سطوح آلودگی گزارش شده توسط شرکت کنترل کیفیت هوای تهران همخوانی دارد. جدول ۱۰ گزارش رسمی کیفیت هوا برای همان تاریخ را ارائه می‌دهد (Air Quality Control Company, ۲۰۱۷).

با ارزیابی نتایج و در نظر گرفتن مدل‌های به دست آمده، نکات زیر حاصل می‌شوند:

- علی‌رغم پیچیدگی ذاتی داده‌های ماهواره سنتینل-۲، این داده‌ها دقت بالاتری در تخمین ذرات معلق نسبت به ماهواره لندست-۸ نشان می‌دهند.
- بر اساس یافته‌های این مطالعه و همچنین تحقیقات پیشین، استفاده از تصاویر با دقت مکانی بالاتر برای تخمین آلودگی هوا مؤثرتر است، هر چند که این امر با پیچیدگی محاسباتی بیشتری همراه است.
- بالاترین دقت در تخمین $PM_{2.5}$ با استفاده از داده‌های سنتینل-۲ با $r = 0.84$ به دست آمد. با این حال، با ساده‌سازی مدل و حذف باند چهارم، دقت کمی پایین‌تر اما همچنان قابل قبول به میزان $r = 0.82$ حاصل شد. به همین ترتیب، بالاترین دقت برای تخمین $PM_{2.5}$ با استفاده از داده‌های لندست-۸ نیز $r = 0.84$ بود. مدل ساده‌شده با حذف باند چهارم دقتی کمی پایین‌تر و به میزان $r = 0.83$ به دست آورد. جدول ۱۱ مقایسه دقیقی از دقت به دست آمده از داده‌های آموزشی و آزمون برای این چهار مدل ارائه می‌دهد.
- استفاده از الگوریتم‌های یادگیری ماشین و رگرسیون مبتنی بر هسته منجر به ایجاد مدل‌های خطی با دقت بالا شد که تأکید می‌کند بازتاب اتمسفری ارتباط قوی‌تری با باندهای طیفی با طول موج کوتاه‌تر دارد.
- تحلیل رگرسیون غیرخطی، به ویژه با استفاده از مدل‌های توان، نشان داد که بازتاب اتمسفری به طور کلی رابطه خطی خود را با ذرات معلق حفظ می‌کند، حتی زمانی که با روش‌های پیچیده‌تر تحلیل رگرسیون انجام شود.
- روش ارائه شده در این مطالعه امکان تولید نقشه‌های دقیق آلودگی هوا با دقت مکانی بالا را فراهم می‌کند که بهبود قابل توجهی را در مقایسه با تحقیقات قبلی نشان می‌دهد. دقت مکانی نقشه‌های آلودگی تهیه شده در این روش 10×10 متر برای سنتینل-۲ و 30×30 متر برای لندست-۸ می‌باشد دقت در پایش و ارزیابی آلودگی را افزایش می‌دهد.



No data	
Clean Air (0-50)	
healthy air (50-100)	
Unhealthy for sensitive groups (100-150)	
Unhealthy (150-200)	
Dangerous (>200)	

(الف)

(ب)

شکل ۱۰. (الف) تصویر ماهواره‌ای لندست-۸ همراه با ایستگاه‌های ثبت زمینی سطح آلودگی. بر اساس گزارش‌های شرکت کنترل کیفیت هوای تهران، مناطق جنوبی و جنوب شرقی تهران در تاریخ ۲۹ اکتبر ۲۰۱۷ به شدت آلوده بودند. و (ب) تصویری از ماهواره لندست-۸ در همان روز است که معادله ردیف نهم جدول ۷ به بازتاب اتمسفری همین تصویر اعمال شده است. این نقشه کاملاً با میزان آلودگی گزارش‌شده توسط شرکت کنترل کیفیت هوای تهران مطابقت دارد.

جدول ۱۰. جدول شاخص کیفیت هوا (AQI) برای هر ایستگاه در تهران در تاریخ ۲۹ اکتبر ۲۰۱۷، ساعت ۱۱ (نزدیک‌ترین زمان به زمان تصویربرداری ماهواره)

شماره	ایستگاه	O_3	NO_2	SO_2	CO	$PM_{2.5}$	PM_{10}	AQI
۱	Average stations	۴۵	۱۰۵	۲۳	۱۰۹	۱۳۷	۱۱۷	۱۳۷
۲	Max station	۳۲	۸۰	۱۲	۵۹	۱۱۳	۸۶	۱۱۳
۳	Area ۱۴	Non	۵۸	۱۵	۵۸	Non	۷۵	۷۵
۴	Area ۱	۳۷	Non	Non	۳۰	۹۹	۷۴	۹۹
۵	Poonak	Non	Non	Non	Non	Non	Non	Non
۶	Area ۱۳	۳۷	۵۷	۱۹	۳۸	۱۳۷	۸۹	۱۳۷
۷	Area ۶	۳۴	۷۴	۱۳	۶۸	۱۱۷	۹۲	۱۱۷
۸	Area ۷	۳۴	۷۲	۹	۷۱	۱۱۱	۷۳	۱۱۱
۹	Area ۱۸	۲۹	۱۰۵	۱۳	۵۳	۱۳۰	۱۰۴	۱۳۰
۱۰	Area ۲	۲۷	۱۰۰	۲۳	۱۰۰	۱۳۴	۹۹	۱۳۴
۱۱	Area ۲۰	۲۷	۱۰۵	۱۱	۶۴	۱۲۸	۱۱۷	۱۲۸
۱۲	Area ۲	۳۳	Non	Non	۲۷	۸۹	۷۰	۸۹
۱۳	Area ۴	Non	Non	Non	Non	Non	Non	Non
۱۴	Area ۱۰	Non	Non	Non	Non	Non	Non	Non
۱۵	Area ۱۱	Non	Non	۶	۱۰۹	Non	Non	۱۰۹
۱۶	Area ۱۶	Non	۱۰۴	Non	۶۰	Non	Non	۶۰
۱۷	Area ۱۹	۱۵	Non	۸	۵۳	Non	۵۷	۵۷
۱۸	Area ۲۱	۳۵	۶۵	۱۷	۶۰	Non	۹۷	۹۷
۱۹	Area ۲۲	Non	Non	Non	Non	Non	Non	Non
۲۰	Area ۸	۳۶	۷۰	۸	۳۵	۱۰۵	۷۹	۱۰۵
۲۱	Area ۱۵	۳۳	۵۷	۷	۲۸	۹۷	۷۸	۹۷

جدول ۱۱. دقت به دست آمده از داده های آزمون و آموزش برای بهترین روابط

ماهواره	روابط با بالاتری دقت	مدل	R ^۲ (%)	r	RMSE	NRMSE
لندست-۸	$PM_{۲,۵} = -۰.۳۲۱ \times R(۱) + ۱.۰۲۸ \times R(۲) + ۳.۷۸ \times R(۳) - ۴.۳۵ \times R(۴) + ۲.۵$	Train	۷۰,۵۶	۰,۸۴	۷,۰۱	۰,۰۶۸
		Test	۶۷,۲۴	۰,۸۲	۷,۴۸	۰,۰۷۱
	$PM_{۲,۵} = ۸.۰ \times R(۱) + ۵.۵۹ \times R(۲) + ۵.۳$	Train	۶۸,۸۹	۰,۸۳	۷,۲۰	۰,۰۶۹
		Test	۶۴,۰۰	۰,۸۰	۷,۷۱	۰,۰۷۳
سنتینل-۲	$PM_{۲,۵} = ۰,۰۴۱ \times R(۲) + ۰,۰۱۲ \times R(۳) - ۰,۰۰۷ \times R(۴) - ۲,۲$	Train	۷۰,۵۶	۰,۸۴	۶,۹۳	۰,۰۹۹
		Test	۶۸,۸۹	۰,۸۳	۷,۳۲	۰,۱۰۴
	$PM_{۲,۵} = ۰,۰۳۴۹ \times R(۲) + ۰,۰۱۰۱ \times R(۳) - ۱,۱$	Train	۶۷,۲۴	۰,۸۲	۷,۲۷	۰,۱۰۳
		Test	۶۵,۵۱	۰,۸۱	۷,۸۱	۰,۱۱۱

۴- نتیجه گیری

هدف این مطالعه پایش و مدل سازی ذرات معلق در جو به عنوان یکی از خطرناک ترین آلاینده ها شناخته شده، با استفاده از تصاویر چندطیفی ماهواره های لندست-۸ و سنتینل-۲ در منطقه تهران بود. داده های آلودگی هوا برای ماهواره لندست-۸ در بازه زمانی چهار ساله و برای ماهواره سنتینل-۲ به مدت ۱۸ ماه جمع آوری شد. بازتاب اتمسفری که از تصاویر ماهواره ای استخراج شد، به عنوان یکی از ویژگی های کلیدی برای شناسایی آلودگی هوا شناسایی گردید. در این تحقیق از سه روش تحلیل رگرسیون استفاده شد. نتایج نشان داد که رگرسیون خطی با استفاده از باندهای دوم، سوم و چهارم سنتینل-۲ دقیق ترین نتایج را برای پیش بینی ذرات معلق ارائه داده است. همچنین روش SVR-GA (رگرسیون بردار پشتیبانی با الگوریتم ژنتیک) که به عنوان یک روش پیشرفته یادگیری ماشین استفاده شده است، تأیید کرد که باندهای طول موج کوتاه تر بیشترین تأثیر را از ذرات معلق می پذیرند. تحلیل رگرسیون غیرخطی نیز نشان داد که رابطه بین ذرات معلق و تصاویر ماهواره ای تقریباً خطی باقی می ماند. در این تحقیق از داده های مربوط به ماه های مختلف سال استفاده گردید و مدل های توسعه یافته برای بازه های زمانی متنوع آموزش دیدند. ارزیابی های انجام شده نشان داد که مدل GA-SVR در بازه های زمستانی و تابستانی عملکرد پایداری داشته و تغییرات دقت آن بین ماه ها در محدوده ای کمتر از ۱۵٪ قرار دارد. این مطالعه نشان داد که تصاویر سنجش از دور چندطیفی پتانسیل بالایی برای شناسایی آلودگی هوا دارند و مدل های آلودگی هوا با دقت مکانی ۱۰ متر برای تصاویر سنتینل-۲ هر ۵ روز و ۳۰ متر برای تصاویر لندست-۸ هر ۱۶ روز (با ترکیب لندست-۹ هر ۸ روز) ارائه می کنند.

منابع

- Air Quality Control Company. (۲۰۱۷). <https://airnow.tehran.ir/home/DaiyAQIArchive.aspx>
- Ashmeet Singh, N., Harsshit Goenka, N., Prakhar Sahu, N., Venkatesh L., N., Dr. Sindhu Sree M., N., Dr. Pavithra G., N., & Dr. T.C.Manjunath, N. (۲۰۲۳). Design & Development of a Air Pollution Monitoring Using Smartphone. *International Journal of Engineering Technology and Management Sciences*, 7(۳), ۴۸۹-۴۹۳. <https://doi.org/۱۰.۴۶۶۴۷/ijetms.۲۰۲۳.v۰۷i۰۳.۶۸>

- Brueckner, M., & Pforr, C. (۲۰۱۰). Global environmental issues. In *Theory and practice of corporate social responsibility* (pp. ۷۳–۸۸). Springer.
- Chang, C.-C., & Lin, C.-J. (۲۰۱۱). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(۳), ۱–۲۷.
- Fraser, R. S., Kaufman, Y. J., & Mahoney, R. L. (۱۹۸۴). Satellite measurements of aerosol mass and transport. *Atmospheric Environment (1967)*, 18(۱۲), ۲۵۷۷–۲۵۸۴.
- Ganjehi, S., Omidvar, B., Malekmohammadi, B., & Norouzi Khatiri, K. (۲۰۱۷). ASSESSMENT AND DEVELOPMENT OF EMERGENCY TRANSPORTATION INDICATORS (CASE STUDY: INFRASTRUCTURES OF TEHRAN MUNICIPALITY, DISTRICT NO. ۱). *Sharif Journal of Civil Engineering*, 33(۳), ۱۱۷–۱۲۵.
- Gassó, S., & Hegg, D. A. (۲۰۰۳). On the retrieval of columnar aerosol mass and CCN concentration by MODIS. *Journal of Geophysical Research: Atmospheres*, 108(D۱), AAC-۶.
- Glencross, D. A., Ho, T.-R., Camina, N., Hawrylowicz, C. M., & Pfeffer, P. E. (۲۰۲۰). Air pollution and its effects on the immune system. *Free Radical Biology and Medicine*, 151, ۵۶–۶۸.
- Gupta, P., Christopher, S. A., Wang, J., Gehrig, R., Lee, Y. C., & Kumar, N. (۲۰۰۶). Satellite remote sensing of particulate matter and air quality assessment over global cities. *Atmospheric Environment*, 40(۳۰), ۵۸۸۰–۵۸۹۲.
- Gupta, P., & Christopher, S. A. (۲۰۰۹). Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: Multiple regression approach. *Journal of Geophysical Research: Atmospheres*, ۱۱۴(D۱۴).
- Harbula, J., Tuček, P., & Pechanec, V. (۲۰۱۰). Dependence of PM_{۱۰} particles concentration on aerosol optical thickness value from the MODIS data. *Proc. Symposium GIS Ostrava*.
- Huang, X., Maier, A., Hornegger, J., & Suykens, J. A. K. (۲۰۱۷). Indefinite kernels in least squares support vector machines and principal component analysis. *Applied and Computational Harmonic Analysis*, 43(۱), ۱۶۲–۱۷۲.
- Kaufman, Y. J., Tanré, D., Remer, L. A., Vermote, E. F., Chu, A., & Holben, B. N. (۱۹۹۷). Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer (Paper ۹۶JD-۳۹۸۸). *JOURNAL OF GEOPHYSICAL RESEARCH-ALL SERIES-*, 102, ۱۷–۵۱.
- Kumar, A., Yadav, A. K., & Tiwari, S. (۲۰۲۰). Estimating ground-level PM_{۲.۵} concentrations using satellite-based aerosol optical depth and meteorological parameters over Delhi, India. *Environmental Science and Pollution Research*, ۲۷, ۱۳۷۰۰–۱۳۷۱۵.
- Lim, H. S., MatJaffri, M. Z., Abdullah, K., Saleh, N. M., & AlSultan, S. (۲۰۰۴). Remote sensing of PM_{۱۰} from LANDSAT TM imagery. *Acrs, 2004*, ۷۳۹–۷۴۴.
- Liu, Y., Sarnat, J. A., Kilaru, V., Jacob, D. J., & Koutrakis, P. (۲۰۰۵). Estimating ground-level PM_{۲.۵} in the eastern United States using satellite remote sensing. *Environmental Science & Technology*, ۳۹(۹), ۳۲۶۹–۳۲۷۸.
- Loveland, T. R., & Irons, J. R. (۲۰۱۶). Landsat ۸: The plans, the reality, and the legacy. *Remote Sensing of Environment*, 185, ۱–۶.
- Mogaraju, J. K. (۲۰۲۳). *Evaluation of spatial and temporal anomalies of major air pollutants using TROPOMI and MODIS data over the Annamayya district of India*.
- Mozafari, S. M., Hasanlou, M., & Arefi, H. (۲۰۱۹). Air pollution estimation using aerosol optical thickness by oli images in Tehran. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, ۷۷۹–۷۸۲.

- Nasir, A., Chaudhry, A. G., Khalid, Z., & Jabbar, A. (۲۰۱۵). An anthropological investigation of urban life and health hazard. *Science International (Lahore)*, 27, ۶۴۷–۶۴۹.
- Nguyen, N. H., & Tran, V. A. (۲۰۱۴). *Estimation of PM10 from AOT of satellite Landsat image over Hanoi city*.
- Pope Iii, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., & Thurston, G. D. (۲۰۰۲). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Jama*, 287(۹), ۱۱۳۲–۱۱۴۱.
- Retalis, A., & Sifakis, N. (۲۰۱۰). Urban aerosol mapping over Athens using the differential textural analysis (DTA) algorithm on MERIS-ENVISAT data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(۱), ۱۷–۲۵.
- Saraswat, I., Mishra, R. K., & Kumar, A. (۲۰۱۷). Estimation of PM_{۱۰} concentration from Landsat ۸ OLI satellite imagery over Delhi, India. *Remote Sensing Applications: Society and Environment*, ۸, ۱۸۰–۱۹۰.
- Saitta, L. (۱۹۹۵). *Support-Vector Networks SVM*. pdf, 297, 273–297.
- Saleh, S. A. H. (۲۰۱۱). Air quality over Baghdad City using earth observation and Landsat thermal data. *Journal of Asian Scientific Research*, 1(۶), ۲۹۱.
- Sameen, M. I., Al Kubaisy, M. A., Nahhas, F. H., Ali, A. A., Othman, N., & Hason, M. M. (۲۰۱۴). Sulfur dioxide (SO_۲) monitoring over Kirkuk City using remote sensing data. *Journal of Civil & Environmental Engineering*, 4(۵), ۱.
- Saraswat, I., Mishra, R. K., & Kumar, A. (۲۰۱۷). Estimation of PM_{۱۰} concentration from Landsat ۸ OLI satellite imagery over Delhi, India. *Remote Sensing Applications: Society and Environment*, 8, ۲۵۱–۲۵۷.
- Silva, R. A., West, J. J., Zhang, Y., Anenberg, S. C., Lamarque, J.-F., Shindell, D. T., Collins, W. J., Dalsoren, S., Faluvegi, G., & Folberth, G. (۲۰۱۳). Global premature mortality due to anthropogenic outdoor air pollution and the contribution of past climate change. *Environmental Research Letters*, 8(۳), ۰۳۴۰۰۵.
- Sirsikar, S., & Karemore, P. (۲۰۱۵). Review paper on air pollution monitoring system. *International Journal of Advanced Research in Computer and Communication Engineering*, 4(۱), ۲۱۸–۲۲۰.
- Sun, L., Wei, J., Bilal, M., Tian, X., Jia, C., Guo, Y., & Mi, X. (۲۰۱۵). Aerosol optical depth retrieval over bright areas using Landsat ۸ OLI images. *Remote Sensing*, 8(۱), ۲۳.
- Thorsen, T. J., Winker, D. M., & Ferrare, R. A. (۲۰۲۱). Uncertainty in observational estimates of the aerosol direct radiative effect and forcing. *Journal of Climate*, 34(۱), ۱۹۵–۲۱۴.
- Uğuz, H. (۲۰۱۱). A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm. *Knowledge-Based Systems*, 24(۷), ۱۰۲۴–۱۰۳۲.
- Uysal, A. K., & Gunal, S. (۲۰۱۲). A novel probabilistic feature selection method for text classification. *Knowledge-Based Systems*, 36, ۲۲۶–۲۳۵.
- Wang, J., & Christopher, S. A. (۲۰۰۳). Intercomparison between satellite-derived aerosol optical thickness and PM_{۲.۵} mass: Implications for air quality studies. *Geophysical Research Letters*, 30(۲۱).
- Yu, X., Nichol, J., Lee, K. H., Li, J., & Wong, M. S. (۲۰۲۲). Analysis of long-term aerosol optical properties combining AERONET sunphotometer and satellite-based observations in Hong Kong. *Remote Sensing*, 14(۲۰), ۵۲۲۰.
- Zhong, Z. (۲۰۱۳). *Proceedings of the International Conference on Information Engineering and Applications (IEA) 2012*. Springer.

Zou, B., You, J., Lin, Y., Duan, X., Zhao, X., Fang, X., Campen, M. J., & Li, S. (۲۰۱۹). Air pollution intervention and life-saving effect in China. *Environment International*, 125, ۵۲۹–۵۴۱.

نسخه پیش انتشار